

Figure 1. Genetic and physical map of recombinant plasmids pHUR849, pDAL201B, pDAL210B, and pDAL200A. The locations of the *pap* genes shown as open bars. Restriction site abbreviations: A, Apa I; B, Bam HI; C, Cla I; E, Eco RI; K, Kpn I; Sa, Sal I; and S, Sma I.

Q

Figure 2. DNA sequences of pHUR849 (A), pDAL201B (B), pDAL210B (C), and pDAL200A (D) papH structural genes. The nontranscribed DNA strand for each clone is shown. Numbering is from the 5' end. The deduced amino acid sequence for the correct frame is shown below each DNA strand. The first amino acid of the mature protein is +1. Stop codon is marked with asterisks. The restriction sites for *Cla* I (ATCGTA), and *Sma* I (CCCGGG), are underlined once, respectively.

Title: IMMUNOGENIC PILI
PRESENTING FOREIGN PEPTIDES,
THEIR PRODUCTION AND USE
Inventor(s): Peter O'Hanley et al.
DOCKET NO.: 050939/0104

pHUR849	1	ATGAGACTGCGATTCTCTGTTCCACTTTTCTTTTGGCTGTGTGTTGTTTCATGGTGT	60
pDAL200A			
pDAL201B		t	
pDAL210B		t	
	61	67	120
pHUR849		TTTGCCGGTCCGTTTCTCCGCCCGCATGTCCCTTCTGAACTACTGGGAGAAGAGCAC	
pDAL200A		t	a
pDAL201B		t	a
pDAL210B			a
	121		180
pHUR849		GTATGGTGGGACGGCAGGGCTGCTTTTCATGGTGAGGTGTGACACCTGCCTGTACTCTG	
pDAL200A			
pDAL201B			
pDAL210B			
	181		240
pHUR849		GCGATGGAAGACGCTGGCAGATTATTGATATGGGGAAACCCCGTAACGGGATTACAG	
pDAL200A		c	t
pDAL201B		c	t
pDAL210B		c	t
	241		300
pHUR849		ATTGTTTCTCCGACCTGAAAGAAAAATCAGCCTCCGGCTCAGGAATTGTGAATTAAAC	
pDAL200A			c
pDAL201B	t		c
pDAL210B	t		c
	301		360
pHUR849		AGTCAGGGTGGGAACTTTTCTCTGATTCCCGGATAAGGGTGACTTTCGATGGCGTCCGG	
pDAL200A			
pDAL201B			
pDAL210B			t
	361		420
pHUR849		GGTGAACCCCGGATAAGTTTAAATTTATCCGGTCAAGCAAAAGGCATTAATCTGCAGATA	
pDAL200A			a
pDAL201B			
pDAL210B			
	421		480
pHUR849		GCTGATGTCAGGGGAAATATTGCCCGGGCAGGAAAAGTAATGCTGCAATACCATTTGACG	
pDAL200A		c	g
pDAL201B		c	g
pDAL210B		c	g
	481		540
pHUR849		GGTAATGAAGAAGCGCTGGATTACACCTCAGAAATTGTGAGAAACGAAAAAACTTGAA	
pDAL200A			c
pDAL201B			c
pDAL210B			c
	541		588
pHUR849		GCCGGAAATTATTTTGTCTGTCTGGATTCCGGGTCGATTATGAGTGA	
pDAL200A		c	
pDAL201B		c	
pDAL210B		c	

Figure. 3. Comparison of deduced nucleotide sequence of *papH* genes pHUR849, pDAL200A, pDAL201B and, pDAL210B. The nucleotide(nt) identities, compared with the deduced sequence of *PapH* nt sequence *papH* gene of pHUR849 (upper case), are indicated by blank space, nt differences for the *PapH* genes of pDAL201B, pDAL210B and, pDAL200A, are shown with the corresponding single letter nt code (lower case), respectively. Numbering is from the 5' end. The first nt of coding for the leader sequence is numbered 1, and the first nt coding for the mature protein is numbered 67.

Title: IMMUNOGENIC PILI
PRESENTING FOREIGN PEPTIDES,
THEIR PRODUCTION AND USE
Inventor(s): Peter O'Hanley et al.
DOCKET NO.: 050939/0104

```

-22                               -1 +1                               18
MRLRFSVPLFFFGCVFVHGVFAGPFPFPPGMSLPEYWGEEH
      C
19                               58
VWWDGRAAFHGEVVRPACTLAMEDAWQIIDMGETPVRDLQ
59                               98
NGFSQPERKFSLRLRNCEFN SQGGNLFSDSRIRVTFDGVR
99                               138
GETPDKFNLSCQAKGINLQIADVRGN IARAGKVMPAIPLT
      a
139                               173
GNEEALDYTLRIVRNGKKLEAGNYFAVLGFRVDYE

```

Figure. 4. Comparison of deduced amino acid sequence of *papH* genes pHUR849, pDAL201B, pDAL210B and, pDAL200A. The vertical arrow indicates the postulated cleavage site for the signal peptidase. The amino acid (aa) identities, compared with the deduced sequence of PapH protein of pHUR849 (upper case), are indicated by blank space, aa differences for the PapH proteins of pDAL201B, pDAL210B and pDAL200A, are shown with the corresponding single letter aa code(lower case), respectively. The first aa of the leader sequence is numbered -22, and the first aa of the mature protein is numbered +1.

Title: IMMUNOGENIC PILI
PRESENTING FOREIGN PEPTIDES,
THEIR PRODUCTION AND USE

Inventor(s): Peter O'Hanley et al.

DOCKET NO.: 050939/0104

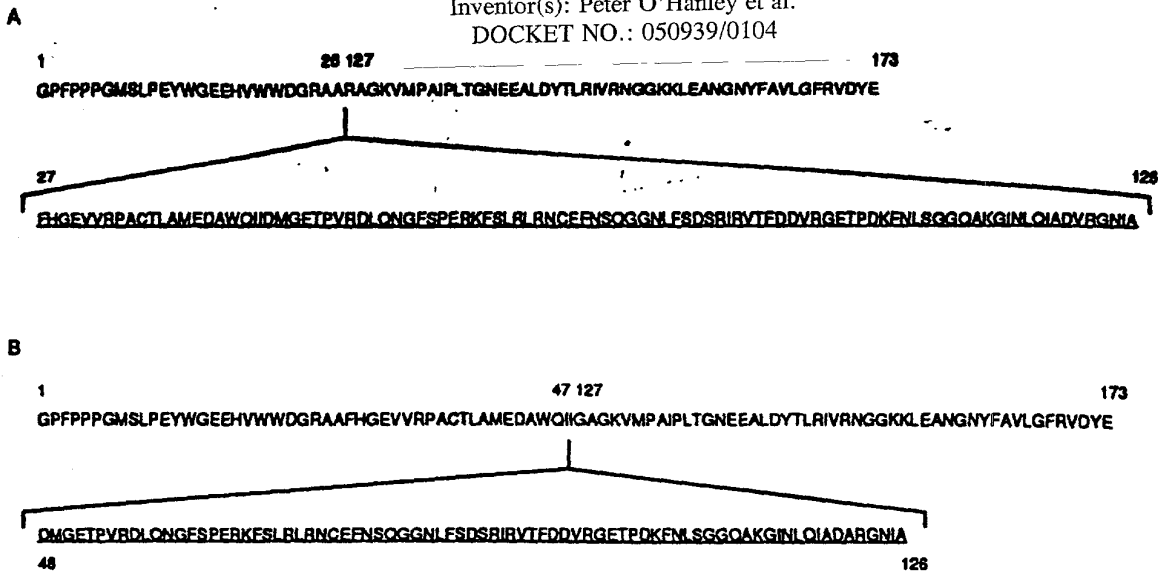


Figure 5. Comparison of deduced amino acid sequence of *papH* gene deletion mutants pHUR849-5(*pap-5*), pDAL201B(*pap-21*), pDAL210B(*pap-17*) and, pDAL200A(*pap-200A*). The deduced amino acid sequence of each of the final constructs is shown, (A) pHUR849-5 and, (B) pDAL201B, pDAL210B and, pDAL200A which are identical to each other. The amino acid identities of the proteins are upper case letters. The first amino acid of the mature fusion protein is numbered 1. The underlined sequence indicates the amino acid residues deleted from the mature fusion protein of each strain.